FIG.2A

```
LOCUS       HSU22027     7215 bp      DNA          PRI       22-OCT-1995
DEFINITION  Human cytochrome P450  (CYP2A6V2)  gene, complete cds.
ACCESSION   U22027
NID         g1008461
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 7215)
  AUTHORS   Fernandez-Salguero,P., Hoffman,S.M., Cholerton,S.,Mohrenweiser,H.,
            Raunio,H., Rautio,A., Pelkonen,O., Huang, J.D., Evans, W.E.,
            Idle, J.R. et,al.
  TITLE     A genetic polymorphism in coumarin 7-hydrozylation: sequence of the
            human CYP2A genes and identification of variant CYP2A6 alleles
  JOURNAL   Am. J. Hum. Genet. 57 (3), 651-660 (1995)
  MEDLINE   95397851
REFERENCE   2 (bases 1 to 7215)
  AUTHORS   Fernandez-Salguero,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAR-1995)  Pedro Fernandez-Salguero,  National
            Institutes of Health, 9000 Rockville Pike, Bethesda, MD 20894, USA
FEATURES              Location/Qualifiers
  source              1..7215
                      /organism="Homo sapiens"
```

FIG.2A CONT.

```
5'UTR           782..790
CDS             join (791..970, 1237..1399, 2115..2264, 2499..2659,
                3207..3383, 4257..4398,4873..5060,5577..5718, 6308..6489)
                /gene=CYP2A6V2:
                /codon_start=1
                /product=cytochrome P450"
                /db_xref-PID:g1008462"
                /translation=MLASGMLLVALLACLTVMVLMSVWQQRKSKGKLPPGPTPLPFIG
                NYLQLNTEQMYNSLMKISERYGPVFTIHLGPRRVVVLCGHDAVREALVDQAEEFSGRG
                EQATFDWVFKGYGVVFSNGERAKQLLRFAIATLRDFGVGKRGIEERIQEESGFLIEAI
                RSTHGANIDPTFFLSRTVSNVISSIVFGDRFDYKDKEFLSLLRMMLGIFQFTSTSTGQ
                LYEMFSSVMKHLPGPQQQAFQLLQGLEDFIAKKVEHNQRTLDPNSPRDFIDSFLIRMQ
                EEEKNPNTEFYLKNLMMSTLNLFIAGTETVSTTLGYGFLLLMKHPEVEAKVHEEIDRV
                IGKNRQPKFEDRAKMPYMEAVIHEIQRFGDVIPMSLARRVKKDTKFRDFFLPKGIEVF
                PMLGSVLRDLRFFSNPRDFNPQHFLGEKGQFKKRDAFVPFSIRKRNCFGEGLARMELF
                LFFTTVMQNFRLKSSQSPKDIDVSPKHVGFATIPRNYTMSFLPR
exon            791..970
                /gene=CYP2A6V2:
                /number=1
exon            1237..1399
                /gene=CYP2A6V2:
                /number=2
exon            2115..2264
```

FIG.2A CONT.

```
                /gene=CYP2A6V2:
                /number=3
     exon       2499..2659
                /gene=CYP2A6V2:
                /number=4
     exon       3207..3383
                /gene=CYP2A6V2:
                /number=5
     exon       4256..4398
                /gene=CYP2A6V2:
                /number=6
     exon       4873..5060
                /gene=CYP2A6V2:
                /number=7
     exon       5577..5718
                /gene=CYP2A6V2:
                /number=8
     exon        6308..6489
                /gene=CYP2A6V2:
                /number=9
     3'UTR      6490..6744
BASE COUNT      1646 a   2196 c   1746 g   1627 t
ORIGIN
```

FIG.2A CONT.

BASE COUNT

```
   1 aagttcccct gaaatatggc tctggtcttc ctccccttgc caatgaagaa gatggcagtg
  61 gaggttctat ggcagccatc ctggcctcac tctgaggttc caatgaggat tctgggcatc
 121 aagagacagc tctgggcaaa gctaaatcaa gtcagcccct ggacccagtg ctgggctgct
 181 gggctttctg ggagaacgcc gctgggcttg ctacacactc ctcctcccag aaactccaca
 241 cccacagccc tgggtcttcc tagccccgag actttcaagt ccatatgcct ggaatccccc
 301 ttcctgagac ccttaaccct gcatcctcca caacagaaga cccctaaatg cacagccaca
 361 ctttgtctta ccctaataaa acccagacct ttggattcct ctcccctgga acccccagat
 421 ccgcacaact ttggggtgca ttctcactct cagaccccaa atccaaagcc caagtgctcc
 481 cctatgcaaa tattccaaac tcctcagttc tacagcttat ctgttgcccc ctcctaaatc
 541 cacagccctg cggcaccctt cctgaagtac cacagattta gtctggaggc cccctctctg
 601 ttcagctgcc ctgggggtccc cttatcctcc cttgctggct gtgtcccaag ctaggcagga
 661 ttcatggtgg ggcatgtagt tgggaggtga aatgaggtaa ttatgtaatc agccaaagtc
 721 catccctctt ttcaggcag tataaaggca aaccacccca gccgtcacca tctatcatcc
 781 ctctaccacc atgctggcct cagggatgct tctggtggcc ttgctggcct gcctgactgt
 841 gatggtcttg atgtctgttt ggcagcagag gaagagcaag gggaagctgc ctccgggacc
 901 caccccattg cccttcattg gaaactacct gcagctgaac acagagcaga tgtacaactc
 961 cctcatgaag gtgtcccaag acagggagat gggtgtctcg gggtgggggc tgcctagttg
1021 gctgggggctt tgtggcaggg ggttgaccag tgtggaccag agtcttagga aatggagttt
1081 tggagtttca gcatcagaaa gacaggatct tgggatgtcc agctcccctga ctgtgagaac
1141 ctgggtgcga agcatcccag cacatgacat ctcggtgctg ggcccccattc agagtggagg
1201 gttctccctc taaccactcc cacccacctc catcagatca gtgagcgcta tggcccccgtg
1261 ttcaccattc acttgggggcc ccggcgggtc gtggtgctgt gtggacatga tgccgtcagg
```

FIG.2A CONT.

```
1321 gaggctctgg tggaccaggc tgaggagttc agcgggcgag gcgagcaagc caccttcgac
1381 tgggtcttca aaggctatgg tgcccaagag ggggaaggtg ggcaggtgga cacgaaggtc
1441 tcagtgttcc cagccttctc cctgactctc ctgacaactg gaggataagg gagagtcccc
1501 agtctggtct tccctcccca tctccctaca ttggggcctc tccatgtgta tccctcacct
1561 gtctccagcg gccctgtcct gattcctccc tgcctctctc tgccccacct ccttattctc
1621 tctcactgga gtctcctctt tcccctctct tcccctctct ctccatctct tgggtttctg
1681 tttaccagcc ctgggtctct gtctacatga gtctttgagg ccctcttagc ttctgggctt
1741 ctctgggttt ctcatctctc cggatccctt tctcaattct tcctctgtct taggatgcca
1801 gggttattcc tacttccaca tcttcaggct ccatctcctg gtaacagtct ctcttccttc
1861 cagaccctct ctgtttctat ctcaatatta aactctctgc tccagctcag cttaagaatc
1921 tcacaccaag agaggatgtc ctccacccag atctccccat atctcactac cccaccctcc
1981 atcctctgcc tccatcactc tctttctctc cccactgccc tgcggacgcg atccaatgga
2041 gtgtggagct aatgccgtga agctatgtgc atctctctgt ctggccgtac ctgggtaata
2101 acctgatcga ctaggcgtgg tattcagcaa cggggagcgc gccaagcagc tcctgcgctt
2161 tgccatcgcc accctgaggg acttcggggt gggcaagcga ggcatcgagg agcgcatcca
2221 ggaggagtcg ggcttcctca tcgaggccat ccggagcacg cacggtgagc aggggacccc
2281 gagtgcgggg gcaggagaag gaaaacaccc aggacgagga acccgcgcgc gttctgcctg
2341 gggatgggga ctaggtgggg aaaggcgccc gcacttccag ccctggagtc tggcgctggg
2401 aatttggctc aacaaggccc tgcctcctgg aattctgact ctcctcagac ctctgagttg
2461 actctctccc caacccccct ctcccgacat acccggaggc gccaatatcg atcccacctt
2521 cttcctgagc cgcacagtct ccaatgtcat cagctccatt gtctttgggg accgctttga
2581 ctataaggac aaagagttcc tgtcactgtt gcgcatgatg ctaggaatct tccagttcac
2641 gtcaacctcc acggggcagg taatggttgc agcccggccc gtgaaggccc ttaccaaaac
```

FIG.2A CONT.

```
2701 cggcaaattg ttcccctacc gggggaaggg ggccccaaat tcccaccgcc ccccggacag
2761 tgtcccctca aaatcagtcc ccgatttggg caaattggca gagtggaacc agacccgggt
2821 tggttgtcca atcccctgct ctccagggac accgggatag cacaacagat gctccccaaa
2881 acagagcctg ctggcaggat gcataccctc agctcagctc tctcaccctg ggcacgtgtt
2941 cccatcccca acttaccggt aatttctaac agatgctccc tacccaggtc ttcttgaata
3001 ttttaacacc cggaaaccct gggtacctaa ccttcctgt  aaactttaga gattagttcc
3061 tatccggccc ctctgaaata cctaaccacc ggagaccaga tgcctttaac tcagttcctt
3121 ccttgctatg aaacaaatcc cattcccatc agctcctgcc ccgtgacagc tgtccttccc
3181 ttcccatcct ctctctgcaa ccccagctct atgagatgtt ctcttcggtg atgaaacacc
3241 tgccaggacc gcagcaacag gcctttcagt tgctgcaagg gctggaggac ttcatagcca
3301 agaaggtgga gcacaaccag cgcacgctgg atcccaattc cccacgggac ttcattgact
3361 cctttctcat ccgcatgcag gaggtacacc ccagcagcca ctgcggggag atgcaaagcc
3241 aggcagaggg aaatcagtct gggagtgggg caggcagatg acacaggccc attcaaatta
3481 accctcatca taataatcct cacaattggc tgggtgccgt ggctaacagc ctgtaatccc
3541 agcactttgg gaggccgagg caggtggatc acctgaggtc aggagttcga gaccagcctg
3601 gccaacatgg tcaaaccccg tctctactaa aaatccaaaa attagttggg catggtggcg
3661 cgaaggggggg cagaggttgc aatgagccaa gatcacggca ttgcactcca gtctgggtga
3721 cagaatgagg ccctgtgtca aaaaaaatta atcacttgtt taaaaagtaa gtgagcctgc
3781 atggtcatgc gcatgtgcag ctccagctac tcaggaggct gaggctggag gattgcttga
3841 gctcaggagt tggcgtccgg cctgtgcaac ttagcaagac caagtcagta taagaaaaaa
3901 aaaaaacaaa aaaaaagctg acagctaagt tgataattga cggacagatg gtcagcaagg
3961 taacgaaggt gagaaggaag agcattgggg gcaacgccag gagtcagggc aagggctggt
```

FIG.2A CONT.

```
4021 tcctagagcg agtctggtag gatctagggc ccctcttctc caccctgcgg tcttgcccca
4081 aagagagggtc gagggtgctg ggattgcgct agactcgagt ctgtgtagat cttgggggtcc
4141 cctcttgacc cccattggtc tgaacctaag agtggaagat ccatgggggtg aacccctaga
4201 tggtgcccctg aggtcaagca ggagtgaggt tgtcctaaag ccccctctcc cttcaggagg
4261 agaagaaccc caacacggag ttctacttga agaacctgat gatgagcacg ttgaacctct
4321 tcattgcagg caccgagacg gtcagcacca ccctgcacta tggcttctta ctgctcatga
4381 agcacccaga ggtggagggt aaggctggag ggggacggaa gtggagggcc ccagaccctc
4441 aaaattcccc ttcgactggt gcaatgtccc cacctgtccc agatcccggg accctgagac
4501 gtgacttgct gtccagagac agggcaacat tcagctggta ggcatcagct gagtctcatt
4561 agatattaaa atattgaaaa tgtctgcact gattggtcag tcacttctgt cccaagccca
4621 ctgagtgccc actgcccgtt ccaccgggtc atcccctaag ttccctcccg tgccctccct
4681 gtgattctgg cacaacctgg ttaacaggat cctactccaa caatgcgaat gggtgatgtc
4741 tgttctgtta tgaatgctct acttccgtct cataggcgga ggcatttcat ccacccatt
4801 ttgcctatcc ggactatcat ttcctgctct gagacccctа gatacctaaa cacattcccc
4861 ctcctccccc agccaaggtc catgaggaga ttgacagagt gatcggcaag aaccggcagc
4921 ccaagtttga ggaccgggcc aagatgccct acatggaggc agtgatccac gagatccaaa
4981 gatttggaga cgtgatcccc atgagtttgg cccgcagagt caaaaaggac accaagtttc
5041 gggatttctt cctcccctaag gtgctatccg cccccacccc ccagactacg gggactccag
5101 cccctctctg tgtcccccagc atcccaccca cattagaagc tttctagacc ctgtcccact
5161 ccctcaatca gtcaaaaaag acttcccccaa ccaccacatc cgttccacct ttccacttag
5221 acactcctga gtcctgcatc tctccagact ctttgtgtca ggagaatcaa acacatgttc
5281 ccaaacttcc tatcttaaga aacagaagcc cccttcccat tcggcctttt gtcataggga
5341 cagaaatctc aggtcccccа aactcctgcc tagaaggaca tggacccccat gtctcccaaa
```

FIG.2A CONT.

```
5401 cttcctgttt cagagatgtg aaccttctat cccccaaggt cctccctcag aggtccccaa
5461 ttcccatgcc tgccacttcc cctcaccggg gcaccctagt tccccctcca gcccctgtgt
5521 actctcaaca atccccccaac ccgcctcatc acatacacct tcctcctccc tcccagggca
5581 tagaagtgtt ccctatgttg ggctccgtgc tgagagacct caggttcttc tccaaccccc
5641 gggacttcaa tccccagcac ttcctgggtg agaaggggca gtttaagaag cgtgatgctt
5701 ttgtgccctt ctccatcagt aagagaccac tgtttggtgc caggcttact actcacacca
5761 gcaggggcct cccttaccca gttccccctct ctgccgtgta gcctagtatt tccccagctt
5821 ggcaagttcc tgttagcaat ctaccgtcga gccaccaggt gatactccct taactaccaa
5881 gcacccagta cctgtgccca ggcaaaagga aaggaaacat catacccctt tcagaggcgg
5941 gggaaaacca aaggccagag agaatcagag atttatttcc ctagggtcac acaggagatt
6001 cttcagcatc cctaaaaagg agatgacggc acagcaggtc atatttggga gttcttatct
6061 gggggaaggg ggatcttaaa cctcccattg tggacacctg gcatcgatca accccatctt
6121 ttggtcatct tttgggtcac tcaaggaaac tgaggtcaag gagggtcaag aggctccctc
6181 ttaaagtctc tcagggccat atattccacc cttcctccct gggagagccg cagctggagg
6241 tcggtactgg ggcgaggctg cactgagagt gggcttcacc tccacccctc ccgcctctcc
6301 tcctcaggaa agcggaactg tttcggagaa ggcctggcca gaatggagct ctttctcttc
6361 ttcaccaccg tcatgcagaa cttccgcctc aagtcctccc agtcacctaa ggacattgac
6421 gtgtccccca aacacgtggg ctttgccacg atcccacgaa actacaccat gagcttcctg
6481 ccccgctgag cgagggctgt gccggtgaag gtctggtggg cggggccagg gaaagggcag
6541 ggccaagacc gggcttggga gaggggcgca gctaagactg ggggcaggat ggcggaaagg
6601 aaggggcgtg gtggctagag ggaagagaag aaacagaagc ggctcagttc accttgataa
6661 ggtgcttccg agctgggatg agaggaagga aacccttaca ttatgctatg aagagtagta
```

FIG.2A CONT.

```
6721 ataatagcag ctcttatttc ctgagcacgt accccccgtgt caccttttgtt caaaaaccat
6781 tgcacgctca cctaatttgc cacaaaaccc ccttcgaagg ggcgttcatg cccatttttac
6841 acgtgacaaa actgaggctt agaaagttgt ctctgatgtc tcacaaaaca taagtgccca
6901 gaaaatctgc gaacacagat ctgtgcccat agccttctag acagattctt aaaaagcacc
6961 tattcctcac gcaaaacagt ttagtataga atcacatggc ctgaacatcc ctgtccgggg
7021 gagttcccca gagacctggg gggtggttgc cctgccttca ctgcacacat gcccacactc
7081 tcacctactc aacatgctgt gactacccgg gtgtaatctg tgcttgctac cagataaggc
7141 cactgtagcc cattcagagt cagcccaggg acacaacgag acatgactgg acatacaggg
7201 tcagtccatt aacaa (SEQ ID NO:1)
```

FIG.2B

```
LOCUS       HSP452B6     1415  bp    RNA            PRI          29-MAY-1992
DEFINITION  Human MRNA FOR CYTOCHROME P-450IIVB6.
ACCESSION   X13494
NID         g35206
KEYWORDS    Cytochrome; cytochrome P450IIB6.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1415)
  AUTHORS   Miles, J.S.
TITLE       Direct Submission
  JOURNAL   Submitted (10-NOV-1988) Miles J.S., Imperial Cancer Research Fund,
            Lab of Molecular Phrmacology and Drug Metabolism, Hugh Robson
            Building, George Square, Edinburgh, EH8 9XD
REFERENCE   2 (bases 1 to 1415)
  AUTHORS   Miles, J.S., McLaren, A.Q. and Wolf, C.R.
  TITLE     Alternative splicing in the human cytochrome P450IIB6 gene
            generates a high level of aberrant messages
  JOURNAL   Nucleic Acids Res. 17 (20), 8241-8255 (1989)
  MEDLINE   90045947
COMMENT     The sequence is a compilation of genomic and cDNA clones. **map:
            chromosomal location=19q12-13.2;
            Data kindly reviewed (13-NOV-1989) by Miles, J.S.
FEATURES             Location/Qualifiers
```

FIG.2B CONT.

```
     source          1..1415
                     /organism="Homo sapiens"
      misc-feature   9..110
                     /note=exon 1, partial"
      misc-feature   111..273
                     /note=exon 2"
       misc-feature  274..423
                     /note=exon 3"
       misc-feature  424..584
                     /note=exon 4"
      misc-feature   585..761
                     /note=exon 5"
      misc-feature   762..903
                     /note=exon 6"
      misc-feature   904..1091
                     /note=exon 7"
      misc-feature   1092..1233
                     /note=exon 8"
       misc-feature  1234..1415
                     /note=exon 9", coding region"
BASE COUNT      341 a    430 c    328 g    316 t
ORIGIN
        1 gaattccgcc ctgcacccat gaccgcctcc caccagggcc ccgccctctg  ccccttttgg
```

FIG.2B CONT.

```
  61 gaaaccttct gcagatggat agaagaggcc tactcaaatc ctttctgagg ttccgagaga
 121 aatatgggga cgtcttcacg gtacacctgg gacccaggcc cgtggtcatg ctgtgtggag
 181 tagaggccat acgggaggcc cttgtggaca aggctgaggc cttctctggc cggggaaaaa
 241 tcgccatggt cgacccattc ttccggggat atggtgtgat ctttgccaat ggaaaccgct
 301 ggaaggtgct tcggcgattc tctgtgacca ctatgaggga cttcgggatg ggaaagcgga
 361 gtgtggagga gcggattcag gaggaggctc agtgtctgat agaggagctt cggaaatcca
 421 agggggccct catggacccc accttcctct tccagtccat taccgccaac atcatctgct
 481 ccatcgtctt tggaaaacga ttccactacc aagatcaaga gttcctgaag atgctgaact
 541 tgttctacca gacttttca ctcatcagct ctgtattcgg ccagctgttt gagctcttct
 601 ctggcttctt gaaatacttt cctggggcac acaggcaagt ttacaaaaac ctgcaggaaa
 661 tcaatgctta cattggccac agtgtggaga agcaccgtga aaccctggac cccagcgccc
 721 ccaaggacct catcgacacc tacctgctcc acatggaaaa agagaaatcc aacgcacaca
 781 gtgaattcag ccaccagaac ctcaacctca acacgctctc gctcttcttt gctggcactg
 841 agaccaccag caccactctc cgctacggct tcctgctcat gctcaaataccctcatgttg
 901 cagagagagt ctacagggag attgaacagg tgattggccc acatcgccct ccagagcttc
 961 atgaccgagc caaaatgcca tacacagagg cagtcatcta tgagattcag agattttccg
1021 accttctccc catgggtgtg ccccacattg tcacccaaca caccagcttc cgagggtaca
1081 tcatccccaa ggacacagaa gtatttctca tcctgagcac tgctctccat gacccacact
1141 actttgaaaa accagacgcc ttcaatcctg accactttct ggatgccaat ggggcactga
1201 aaaagactga agcttttatc cccttctcct tagggaagcg gatttgtctt ggtgaaggca
1261 tcgcccgagc ggaattgttc ctcttcttca ccaccatcct ccagaacttc tccatggcca
1321 gccccgtggc cccagaagac atcgatctga caccccagga gtgtggtgtg ggcaaaatac
1381 ccccaacata ccagatccgc ttcctgcccc gctga (SEQ ID NO:2)
```